



**University of  
Zurich<sup>UZH</sup>**

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2016

---

## **Gründe für Vertrauen, Vertrauenswürdigkeit und Kompetenz**

Budnik, Christian

**Abstract:** One reason that trust is interesting for philosophers is that there are reasons for trust. We want to know when it is appropriate to trust other people, because trusting foolishly can be dangerous. The paper argues that there are two competing ways of understanding reasons of trust – a non-voluntarist and a voluntarist way. On the non-voluntarist picture, reasons for trust are associated with the trustworthiness of another person. On the voluntarist picture, you can appropriately trust another person without judging her to be trustworthy. In the paper a case is made against the voluntarist interpretation. At the same time it is argued that non-voluntarist theories of trust are faced with a particular problem. This problem has to do with the assumption that in order to appropriately trust another person you have to judge her competent in the relevant area of interaction. It is argued that this competence assumption should not be understood as a necessary condition for trustworthiness but rather as a necessary condition for mere reliability. Doing so helps to solve problems associated with trust in medicine ethics and political philosophy, and it opens up new perspectives for theories of trust.

DOI: <https://doi.org/10.1515/dzph-2016-0007>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-187280>

Journal Article

Published Version

Originally published at:

Budnik, Christian (2016). Gründe für Vertrauen, Vertrauenswürdigkeit und Kompetenz. Deutsche Zeitschrift für Philosophie, 64(1):103-118.

DOI: <https://doi.org/10.1515/dzph-2016-0007>

Christian Budnik\*

# Gründe für Vertrauen, Vertrauenswürdigkeit und Kompetenz

DOI 10.1515/dzph-2016-0007

**Abstract:** One reason that trust is interesting for philosophers is that there are reasons for trust. We want to know when it is appropriate to trust other people, because trusting foolishly can be dangerous. The paper argues that there are two competing ways of understanding reasons of trust – a non-voluntarist and a voluntarist way. On the non-voluntarist picture, reasons for trust are associated with the trustworthiness of another person. On the voluntarist picture, you can appropriately trust another person without judging her to be trustworthy. In the paper a case is made against the voluntarist interpretation. At the same time it is argued that non-voluntarist theories of trust are faced with a particular problem. This problem has to do with the assumption that in order to appropriately trust another person you have to judge her competent in the relevant area of interaction. It is argued that this competence assumption should not be understood as a necessary condition for trustworthiness but rather as a necessary condition for mere reliability. Doing so helps to solve problems associated with trust in medicine ethics and political philosophy, and it opens up new perspectives for theories of trust.

**Keywords:** trust, trustworthiness, voluntarism, nonvoluntarism, competence

Daß Vertrauen philosophisch interessant ist, liegt unter anderem daran, daß es *Gründe für Vertrauen* gibt. Personen sind rational kritisierbar dafür, daß sie anderen Personen entweder gar nicht oder zu sehr vertrauen. Blindes oder leichtfertiges Vertrauen ist gefährlich und sollte um jeden Preis vermieden werden. Gar nicht zu vertrauen, ist vor dem Hintergrund der Tatsache, daß wir in unzähligen Kontexten auf die Kooperation mit anderen Personen angewiesen sind, keine Option. Was heißt es aber, daß eine Person einen Grund für Vertrauen hat? Wie lassen sich solche Gründe näher charakterisieren? Was zeichnet angemessenes Vertrauen gegenüber blindem Vertrauen aus? Solche Fragen scheinen extrem wichtig zu sein, sie lassen sich allerdings nicht beantworten, ohne im Vorfeld die

---

\*Kontakt: Christian Budnik, Universität, Philosophisches Institut, Länggäßstraße 49a, 3012 Bern, Schweiz, christian.budnik@philo.unibe.ch

Frage danach zu klären, was Vertrauen überhaupt ist: Nur wenn wir wissen, um welch ein Phänomen es sich bei Vertrauen handelt, werden wir in der Lage sein, eine Bestimmung der Gründe für Vertrauen zu formulieren. Vor diesem Hintergrund verwundert es nicht, daß ein Großteil der philosophischen Debatte um die Frage danach kreist, was für ein mentaler Zustand Vertrauen ist.

Für den vorliegenden Kontext ist es besonders aufschlußreich, die in diesem Zusammenhang vertretenen Positionen in zwei Gruppen einzuteilen. Zur ersten Gruppe gehören Theorien, die Vertrauen als kognitive oder affektive Reaktion auf die Vertrauenswürdigkeit einer anderen Person verstehen. Zur zweiten Gruppe gehören Theorien, die die Gründe, die für bzw. gegen Vertrauen sprechen, von der Vertrauenswürdigkeit der Person, in die Vertrauen gesetzt werden soll, abkoppeln. Im Folgenden werde ich zunächst die Opposition dieser beiden Theorietypen, die ich als nicht-voluntaristische bzw. voluntaristische Ansätze bezeichnen werde, rekonstruieren und für die Angemessenheit des von nicht-voluntaristischen Ansätzen vertretenen Zusammenhangs zwischen Gründen für Vertrauen und Vertrauenswürdigkeit plädieren (Abschnitte 1 und 2). In einem zweiten Schritt werde ich auf ein zentrales Merkmal von nicht-voluntaristischen Theorien aufmerksam machen – die Annahme, daß vertrauenswürdige Personen immer auch kompetente Personen sein müssen – und Zweifel an der Angemessenheit dieser Annahme anmelden (Abschnitt 3). Abschließend werde ich andeuten, welche Konsequenzen sich sowohl für die Analyse des Vertrauensbegriffes als auch für einige der angewandten Kontexte, in denen Vertrauen eine wichtige Rolle spielt, ergeben, sobald man die Kompetenzannahme ganz fallen läßt (Abschnitt 4).

## 1 Gründe für Vertrauen und Vertrauenswürdigkeit

Zu den nicht-voluntaristischen gehören jene Ansätze, die Vertrauen entweder mit einer emotionalen bzw. affektiven Einstellung oder mit einer spezifischen Überzeugung identifizieren. Die gemeinsame Idee dahinter läßt sich etwa folgendermaßen auf den Punkt bringen: Eine wahre Überzeugung entspricht dem, was der Fall ist; bei einer angemessenen Emotion muß das formale Objekt dieser Emotion zu ihrem Bezugsobjekt in der Welt passen. Wenn ich denke, daß es draußen regnet, während tatsächlich die Sonne scheint, habe ich auf eine ähnliche Weise versagt wie in dem Fall, in dem ich mich vor einem ausgestopften Spielzeugbären fürchte. Vertrauen – genauer: gerechtfertigtes oder angemessenes Vertrauen – scheint auf ganz ähnliche Weise in einem normativen Verhältnis des ‚Passens‘ zu der Welt zu stehen wie Überzeugungen oder Emotionen, indem es eine passende

Reaktion auf einen bestimmten Sachverhalt in der Welt darstellt: Es ist genau dann angebracht, einer Person zu vertrauen, wenn diese Person vertrauenswürdig ist.<sup>1</sup> Sollte sich herausstellen, daß die Person doch nicht vertrauenswürdig ist, ist es an dem rationalen Akteur, davon Abstand nehmen, dieser Person zu vertrauen. Zumindest auf den ersten Blick handelt es sich bei Vertrauen demnach um eine Einstellung mit einer Geist-zu-Welt-Passensrichtung.<sup>2</sup> Zudem – und hier liegt der eigentliche Kontrast zu den voluntaristischen Ansätzen – kann Vertrauen gemäß dieser Interpretation nicht als etwas verstanden werden, wozu sich eine rationale Person ohne Rücksicht auf das potentielle Vertrauensobjekt *entscheiden* kann – sondern es muß eben die Frage der Vertrauenswürdigkeit berücksichtigt werden.

Die Frage nach Kriterien für Vertrauenswürdigkeit wird im Rahmen nicht-voluntaristischer Ansätze entweder direkt zum Thema gemacht oder indirekt beantwortet, indem der Gehalt der Einstellung, mit der man Vertrauen identifiziert, spezifiziert wird. So versteht z. B. Hardin Vertrauen als die Überzeugung, daß es in dem Interesse der Person, der ich vertraue, liege, meine Interessen zu befördern,<sup>3</sup> wohingegen Baier Vertrauen als die Überzeugung interpretiert, daß die Person, der man vertraut, einem mit Wohlwollen begegne.<sup>4</sup> Beide Positionen gehen demnach davon aus, daß Vertrauen eine spezifische Überzeugung darstellt, aber sie unterscheiden sich in der Bestimmung des Kriteriums für Vertrauenswürdigkeit: Während für Hardin eine vertrauenswürdige Person die Person ist, die meine Interessen in ihren Interessen eingeschlossen hat, versteht Baier sie als diejenige Person, die mir Wohlwollen entgegenbringt. Umgekehrt sind diese spezifischen Interpretationen von Vertrauenswürdigkeit keinesfalls an eine kognitivistische Interpretation von Vertrauen gebunden. Dies sieht man etwa daran, daß Jones Vertrauen mit der affektiven Einstellung des Optimismus identifiziert, den Gehalt dieser Einstellung aber dahingehend interpretiert, daß es sich dabei um eine optimistische Haltung bezüglich des Wohlwollens handelt, das mir die Person, der ich vertraue, entgegenbringt.<sup>5</sup> Damit unterscheidet sie

---

1 Vgl. Jones (2010), 68 ff.

2 Zum Begriff der ‚direction of fit‘ vgl. etwa Platts (1979), 255 ff., oder Humberstone (1992).

3 Vgl. Hardin (2002).

4 Vgl. Baier (1986).

5 Vgl. Jones (1996). Für einen weiteren auf einer affektiven Einstellung basierenden Ansatz vgl. McLeod (2002), wo Vertrauen als optimistische Einstellung bezüglich der Integrität der Person, der man vertraut, verstanden wird. Während alle kognitivistische Ansätze sich so konstruieren lassen, daß sie Vertrauen als eine angemessene Reaktion auf Vertrauenswürdigkeit verstehen, ist dies bei affektbasierten Ansätzen keinesfalls die Regel, d. h., es werden Ansätze vertreten, nach denen Vertrauen zwar eine affektive Einstellung ist, aber nicht als Reaktion auf Vertrauenswürdigkeit verstanden wird; vgl. hierzu Becker (1996).

sich von Baier in der Frage der Bestimmung des mentalen Zustandes, mit dem Vertrauen identifiziert wird, ist mit ihr aber auf einer Linie, was das Kriterium der Vertrauenswürdigkeit betrifft.

Woran läßt sich aber absehen, ob ein Kriterium für Vertrauenswürdigkeit angemessener ist als ein Kriterium, das von einem konkurrierenden Ansatz vorgeschlagen wird? Die wichtigste Strategie besteht in diesem Zusammenhang darin, auf den Unterschied zwischen Vertrauen und bloßem Sich-Verlassen – oder, für den vorliegenden Kontext entscheidend: zwischen Vertrauenswürdigkeit und bloßer Verlässlichkeit – aufmerksam zu machen. In Fällen, in denen ich einer Person vertraue – so die von Baier prominent in die Debatte eingeführte Idee – ist mehr involviert als in Situationen, in denen ich gute Gründe habe, davon auszugehen, daß diese Person etwas Bestimmtes tun wird, weil es z. B. in ihrem Eigeninteresse liegt, weil sie Angst hat, weil sie dazu manipuliert wurde oder aufgrund einer fest verankerten Handlungsdisposition nicht anders als erwartet handeln kann. So konnten sich die Bürger von Königsberg – um Baiers Beispiel zu verwenden – darauf verlassen, daß Kant jeden Morgen um sechs Uhr in der Früh an einer bestimmten Straßenecke auftauchte – dennoch haben sie ihm nicht vertraut, und entsprechend wäre es wiederum schief, Kant als vertrauenswürdig in Hinblick auf die Uhrzeit zu charakterisieren. Ganz analog ließe sich von dieser Warte aus gegen Hardins Kriterium für Vertrauenswürdigkeit argumentieren, daß es ganz verschiedene Weisen gibt, wie die Interessen einer Person in den Interessen einer anderen Person eingeschlossen sein können, und daß nicht alle von ihnen kompatibel mit Vertrauen bzw. Vertrauenswürdigkeit sind. In einer Situation etwa, in der ich denke, daß ein Bekannter mir einen Gefallen tun wird, weil ich weiß, daß er vor mir Angst hat, würde Hardin behaupten müssen, daß die Tatsache, daß er Angst vor mir hat, dafür spricht, ihn zumindest in dem betreffenden Handlungskontext als vertrauenswürdig zu betrachten. Baier scheint aber ganz recht damit zu haben, daß in so einem Fall von Vertrauen nicht die Rede sein kann, und daß es höchstens wahr ist, daß ich mich darauf verlassen kann, daß der Bekannte mir einen Gefallen tun wird.

Ob man den für Vertrauen charakteristischen Aspekt am besten über den Begriff des Wohlwollens zu fassen bekommt – wie Baier und Jones denken – sei an dieser Stelle dahingestellt.<sup>6</sup> Wichtiger ist Baiers Beobachtung, daß die emotionalen Reaktionen auf den Fall, in dem meine Erwartungen nicht erfüllt werden, unterschiedlich ausfallen werden, je nachdem, ob es sich um einen Fall handelt, in dem ich mich auf eine Person lediglich verlassen habe, oder um den Fall, in dem ich der Person vertraut habe: In dem einen Fall werde ich mich ärgern,

---

<sup>6</sup> Vgl. allerdings die durchaus einschlägigen Einwände in Holton (1994), 64–65.

so wie ich mich vielleicht darüber ärgere, daß ein Gerät, auf das ich mich verlassen habe, nicht funktioniert – in dem anderen Fall werde ich mich darüber hinaus *betrogen fühlen*.<sup>7</sup> Der von Baier auf diese Weise eingeführte Unterschied zwischen dem Gefühl der Enttäuschung und dem Gefühl des Betrogenseins als unterschiedliche Reaktionen auf mißlungene Formen des Sich-Verlassens und des Vertrauens ist zu einer Art Lackmustest für Vertrauen geworden: Den Unterschied zwischen beiden Formen der emotionalen Reaktion erklären zu können und damit der Unterscheidung zwischen Vertrauen und bloßem Sich-Verlassen gerecht zu werden, stellt inzwischen ein nahezu allseits anerkanntes Adäquatheitskriterium für Vertrauentheorien dar.

Nicht-voluntaristische Ansätze versuchen, diesem Anspruch gerecht zu werden, indem sie Kriterien für Vertrauenswürdigkeit formulieren, die zur Explikation des angesprochenen Unterschieds in emotionaler Reaktion herangezogen werden können.<sup>8</sup> Gleichzeitig bestimmen sie dadurch aber auch, was als Gründe für Vertrauen aufzufassen ist. Wenn Vertrauenswürdigkeit etwa darin besteht, daß eine Person mir gegenüber wohlwollend eingestellt ist, dann habe ich genau dann einen Grund, einer anderen Person zu vertrauen, wenn ich einen Grund habe, zu denken bzw. zu fühlen, daß eine andere Person mir gegenüber in einem bestimmten Handlungskontext wohlwollend eingestellt ist. Gründe für Vertrauen speisen sich dieser Auffassung zufolge wesentlich aus objektiven Tatsachen, die das potentielle oder tatsächliche Objekt des Vertrauens betreffen. Ob mein Vertrauen angemessen ist oder nicht, ist dementsprechend eine Frage, die in der Hauptsache die Person betrifft, der ich vertraue.

## 2 Gründe für Vertrauen ohne Vertrauenswürdigkeit

Eine radikale Abweichung von diesem Verständnis des Zusammenhangs zwischen den Gründen für Vertrauen und Vertrauenswürdigkeit legen die in jüngster Zeit besonders populären voluntaristischen Ansätze nahe. Ein voluntaristischer

---

<sup>7</sup> Vgl. Baier (1986), 99: „The trusting can be betrayed, or at least let down, and not just disappointed.“

<sup>8</sup> Ein Beispiel zur Illustration: Baier versteht Vertrauen im Wesentlichen als die Überzeugung darüber, daß die Person der vertraut wird, einem gegenüber wohlwollend eingestellt ist. Sollte mein Vertrauen enttäuscht werden, werde ich mich nicht nur ärgern, daß etwas nicht so geklappt hat, wie ich es mir vorgestellt hatte; Baier zufolge führt gerade die Tatsache, daß ich in so einer Situation feststellen muß, daß mir *kein Wohlwollen entgegengebracht wurde*, zu einer tieferen emotionalen Reaktion.

Ansatz unterscheidet sich von einem nicht-voluntaristischen Ansatz dadurch, daß in seinem Rahmen angemessenes Vertrauen als etwas betrachtet wird, wozu man sich jederzeit ohne spezielle Berücksichtigung der jeweils in Frage stehenden Person entscheiden kann. Anders gesagt: Ob mein Vertrauen angemessen ist oder nicht, hängt nicht davon ab, ob die Person, der ich vertraue, auch vertrauenswürdig ist. Der prominenteste Vertreter einer solchen Position ist Richard Holton, der als paradigmatisches Beispiel für Vertrauen die aus Theatergruppen bekannte Situation skizziert, in der eine Person sich nach hinten fallen lassen soll, um von den anderen Teilnehmerinnen aufgefangen zu werden. In so einer Situation, so Holton, liegt es ganz an mir, ob ich mich dazu entschieße, mich fallen zu lassen; und wenn ich mich dazu entschieße, dann heißt das nichts anderes, als daß ich den anderen Teilnehmerinnen vertraue: „If you are like me, there is a moment at which you weigh up whether or not to let yourself fall. How does it feel at that moment? It feels as though you are *deciding* whether or not to trust. I think we should take this feeling at face value: there are circumstances in which we can decide to trust.“<sup>9</sup>

Gleichzeitig wäre es falsch, wollte man behaupten, daß voluntaristische Ansätze keine Antwort auf die eingangs formulierte Frage nach Gründen für Vertrauen liefern würden. Zum einen sind sie in einer minimalen Hinsicht an Vertrauenswürdigkeit interessiert: Um auf angemessene Weise einer Person zu vertrauen, muß ich nämlich in der Annahme gerechtfertigt sein, daß sie *in der Lage ist*, das zu tun, was ich von ihr erwarte. Es wäre dementsprechend unangemessen, sich dazu zu entschließen, den Teilnehmerinnen im ‚Vertrauenskreis‘ zu vertrauen, wenn ich gleichzeitig Grund zu der Annahme hätte, daß sie mit einem Nervengift betäubt sind und mich gar nicht auffangen können. Abgesehen davon ist aber für die Vertreter voluntaristischer Ansätze jede Unsicherheit im Hinblick auf die Motivation der potentiellen Vertrauenspartner kompatibel mit einer Entscheidung, ihnen auf angemessene Weise zu vertrauen. Dies funktioniert sogar in Fällen, in denen ich Grund habe, von einer Person Schlimmstes zu befürchten; solange es lediglich im Bereich des Möglichen ist, daß sie das tun wird, was ich von ihr erwarte, kann ich ihr auf angemessene Weise Vertrauen entgegenbringen.<sup>10</sup>

Vertreter voluntaristischer Ansätze können allerdings noch mehr zu der Frage nach den Gründen für Vertrauen sagen. Dies liegt daran, daß sie Vertrauen

<sup>9</sup> Holton (1994), 63, Hervorh. im Orig.

<sup>10</sup> Vgl. die ‚unpersönliche‘ Analogie ebd., 68: „I can rely on a rope that I doubt to be secure because I have no alternative.“

als eine Form des Sich-Verlassens (‘reliance’) interpretieren.<sup>11</sup> Sich-Verlassen stellt wiederum einen Bestandteil von weitergefaßten Absichten oder Plänen von Akteuren dar.<sup>12</sup> Dies hat zur Folge, daß die Gründe, die für bzw. gegen eine Entscheidung sprechen, sich auf eine andere Person zu verlassen, von der Rolle abhängen werden, die dieses Sich-Verlassen im weiteren Kontext meiner Pläne und Absichten spielt. Ein rationaler Akteur muß demzufolge in jeder Situation eine Abwägung zwischen der Wahrscheinlichkeit, mit der eine andere Person etwas tun wird, und der Relevanz, die dieses potentielle Tun für seine Pläne hat, treffen. Ich habe entsprechend keinen Grund, darauf zu vertrauen, daß eine Person etwas Bestimmtes tun wird, wenn ich denke, daß es sehr unwahrscheinlich ist, daß sie es tun wird, und ich nicht davon abhängig bin, daß sie es tut (z. B. weil ich andere Optionen zur Realisierung meiner Absichten habe). Umgekehrt kann es sein, daß ich einen Grund zum Vertrauen habe, auch wenn ich denke, daß es nicht sehr wahrscheinlich ist, daß eine andere Person etwas tun wird; dies ist dann der Fall, wenn die Realisierung meiner Pläne davon abhängt, daß sie es tut, und wenn die Pläne selbst ein bestimmtes Gewicht für mich haben: Für einen verzweifelten Flüchtling in einem Kriegsgebiet mag es angemessen sein, auf die Hilfe eines Menschenschleusers zu vertrauen, selbst wenn der Flüchtling davon ausgeht, daß es nicht sehr wahrscheinlich ist, daß der Schleuser sein Versprechen halten und ihn heil aus dem Kriegsgebiet herausschaffen wird.

An dieser Stelle der Rekonstruktion der voluntaristischen Position sind eine Ergänzung und eine kritische Bemerkung angebracht. Zunächst zur Ergänzung: Sollten Vertreter voluntaristischer Positionen bei der Ansicht stehen bleiben, daß Vertrauen ein Sich-Verlassen sei, würden sie ziemlich offensichtlich das oben angesprochene Adäquatheitskriterium nicht erfüllen, demzufolge eine Vertrauentheorie den Unterschied zwischen Vertrauen und bloßem Sich-Verlassen explizieren können muß. Dies ist der Grund, weshalb die Identifikation von Vertrauen und Sich-Verlassen von Vertretern voluntaristischer Ansätze durch Komponenten ergänzt wird, die diesem Unterschied und dem damit verbundenen Unterschied in der emotionalen Reaktion auf Fälle, in denen Erwartungen ent-

---

<sup>11</sup> Neben Holton wurden voluntaristische Ansätze zuletzt prominent in Faulkner (2007 u. 2011), sowie Hawley (2012) vertreten. Anstelle von ‘reliance’ spricht Faulkner von ‘dependence’, bleibt in der Sache aber ebenso ein Voluntarist wie Holton. Hawleys Ansatz mag auf den ersten Blick wie ein kognitivistischer aussehen, weil in ihrer Bestimmung von Vertrauen eine spezifische Überzeugung eine wichtige Rolle spielt (vgl. Hawley 2012, 10); tatsächlich vertritt sie aber eine voluntaristische Position, weil der Unterschied zwischen Vertrauen und Mißtrauen auch bei ihr eine Frage der Entscheidung des Akteurs ist.

<sup>12</sup> Vgl. Holton (1994), 72: „When I rely on someone to do something, I work this reliance into my plans: I plan on the supposition that they will do it.“



täuscht werden, Rechnung trägt. Holton etwa bringt in diesem Zusammenhang den an Strawsons Terminologie angelehnten Begriff des Teilnehmerstandpunkts ins Spiel.<sup>13</sup> Dieser Auffassung zufolge vertraue ich genau dann einer Person, wenn ich mich vom Teilnehmerstandpunkt darauf verlasse, daß sie etwas tun wird, wobei der Teilnehmerstandpunkt zum Teil durch die Disposition definiert ist, bestimmte reaktive Einstellungen an den Tag zu legen, sollte meine Erwartung enttäuscht werden.<sup>14</sup>

Wie auch immer man zu den Details dieser Interpretation des Unterschieds zwischen Vertrauen und Sich-Verlassen stehen mag, so stellt sich doch die Frage, wie plausibel die generelle Strategie der voluntaristischen Ansätze ist, die Gründe für Vertrauen von der Vertrauenswürdigkeit der potentiellen Vertrauenspartner abzukoppeln. Obwohl man zur Beantwortung dieser Frage in die Detailanalyse der einzelnen Vorschläge einsteigen müßte,<sup>15</sup> glaube ich, daß eine recht einfache Überlegung die Plausibilität dieser Strategie zumindest in ein zweifelhaftes Licht rückt: Was die Vertreter voluntaristischer Ansätze als Gründe für Vertrauen interpretieren, sind keinesfalls Gründe für Vertrauen, sondern Gründe, etwas zu tun, was so aussieht, als würde man vertrauen. Der Flüchtling, von dem die Vertreter voluntaristischer Theorien behaupten müßten, daß er darauf vertraut, daß ihn der Schlepper aus dem Kriegsgebiet herausschafft, ist tatsächlich sehr weit davon entfernt, dem Schlepper zu vertrauen; seine Situation ist allerdings so dramatisch, daß es rational für ihn ist, ein bestimmtes Risiko einzugehen und sich auf das Angebot des Schleppers einzulassen; für sich betrachtet ist dies eine Handlung, die ununterscheidbar von jener ist, die der Flüchtling ausführen würde, wenn er den Schlepper für vertrauenswürdig hielte – daraus folgt aber keinesfalls, daß sie auch tatsächlich auf Vertrauen basiert. Wenn voluntaristische Ansätze recht hätten – so eine andere Weise, den Kritikpunkt zu formulieren – dann wäre es nicht möglich, sich auf mißtrauische Weise auf andere Personen zu verlassen; dies scheint aber eine recht gewöhnliche und oft sehr angemessene bzw. rational geforderte Weise zu sein, auf die wir mit anderen Personen zu tun haben.

---

13 Vgl. ebd., 66; zu Strawsons Begriff der ‚participant attitude‘ und den ‚reactive attitudes‘ vgl. Strawson (1974).

14 Faulkner operiert in diesem Zusammenhang mit einem normativ aufgeladenen Begriff der Erwartung (vgl. Faulkner 2007, 882), während Hawley die Identifikation von Vertrauen und Sich-Verlassen durch eine Überzeugung über die ‚commitments‘ des potentiellen Vertrauenspartners ergänzt (vgl. Hawley 2012, 10 ff.)

15 Vgl. etwa die ausführliche Kritik in Hieronymi (2008).

### 3 Vertrauenswürdigkeit und Kompetenz

Wenn an den zuletzt formulierten Überlegungen etwas dran ist, dann scheinen voluntaristische Ansätze daran zu scheitern, daß sie den folgenden Zusammenhang falsch konstruieren: Es ist eine Sache, einer anderen Person zu vertrauen, aber eine ganz andere, Handlungen auszuführen, deren Ergebnis zum Teil oder vollständig davon abhängt, wie sich eine andere Person verhalten wird. Für Handlungen können wir ganz verschiedene Gründe haben, und je nachdem, um welche Gründe es sich konkret handelt, werden wir bestimmte Handlungen als nützlich, wertvoll, klug, zielführend oder moralisch richtig auszeichnen können. Wie auch immer man den Begriff des Handlungsgrundes versteht, so scheint gleichermaßen klar zu sein, daß unsere Gründe immer auch Annahmen darüber enthalten werden, was der Fall ist. In einigen Fällen werden dies Annahmen darüber sein, wie sich andere Personen verhalten werden; und in manchen dieser Fälle werden wir diese Annahmen auf der Grundlage von Vertrauen treffen. Hier ist es keinesfalls an uns, unilateral Entscheidungen darüber zu fällen, ob wir einer anderen Person vertrauen oder nicht; ganz im Gegenteil: soll die Tatsache, daß ich einer Person vertraue, auf rationale Weise relevant dafür sein, etwas Bestimmtes zu tun, dann darf dieses Vertrauen nicht willkürlich sein, sondern es muß von der Vertrauenswürdigkeit der Person, um die es in der konkreten Situation geht, abhängen.

Es scheint also, daß nicht-voluntaristische Theorien, die Gründe für Vertrauen in Vertrauenswürdigkeit verorten, voluntaristischen Theorien vorzuziehen sind, die nicht sehr viel zur Frage der Vertrauenswürdigkeit zu sagen haben. Gleichzeitig denke ich, daß nicht-voluntaristische Theorien ihrerseits mit einem grundlegenden Problem konfrontiert sind, für dessen Lösung sich etwas von voluntaristischen Theorien lernen läßt. Um das Problem in den Blick zu bekommen, ist es an dieser Stelle hilfreich, einen Aspekt von nicht-voluntaristischen Theorien zu thematisieren, den ich in meiner bisherigen Rekonstruktion bewußt ausgeklammert habe. Ganz unabhängig davon, ob Vertrauen als eine Überzeugung oder ein affektiver Zustand verstanden wird, machen nicht-voluntaristische Ansätze eine Annahme über Vertrauenswürdigkeit, von der ich glaube, daß sie systematisch nicht unerhebliche Konsequenzen hat, für sich genommen aber extrem problematisch ist. Es handelt sich hierbei um die folgende Kompetenzannahme:

(KA) Eine Person *A* vertraut einer Person *B*, wenn *A* die Überzeugung hat (oder auf affektive Weise optimistisch ist), daß *B* über für den jeweiligen Handlungskontext relevante Kompetenzen verfügt.

(KA) formuliert eine notwendige Bedingung für angemessenes Vertrauen. Gleichzeitig wird durch die Annahme eine notwendige Bedingung für die Vertrauens-

würdigkeit einer Person formuliert. Die Annahme erscheint so selbstverständlich, daß für sie meistens nicht sehr ausführlich argumentiert wird. Sie steht im Hintergrund nahezu aller Positionen, in denen mit dem Begriff der Vertrauenswürdigkeit operiert wird:

Trust, I have claimed, is reliance on others' *competence* and willingness to look after, rather than harm, things one cares about which are entrusted to their care.<sup>16</sup>

At the center of trust is an attitude of optimism about the other person's goodwill. But optimism about goodwill is not sufficient, for some people have very good wills but very little *competence*, and the incompetent deserve our trust almost as little as the malicious. [...] Thus, we should say that trust is optimism about the goodwill and *competence* of another.<sup>17</sup>

First, if A is to trust B, then B must have not only the motivation to do *x* but also the *competence*. An agent who cannot act on my behalf is a poor agent.<sup>18</sup>

Clear conditions for trustworthiness are that the trustworthy person is *competent* and committed to do what s/he is trusted to do. [...] Failing to be optimistic about people's *competence* also makes trust impossible. Without being confident that people will display some *competence*, we cannot trust them.<sup>19</sup>

Inwiefern kann behauptet werden, daß (KA) selbstverständlich erscheint? In diesem Zusammenhang kann an nahezu jeden beliebigen Vertrauenskontext gedacht werden. Angenommen, ich vertraue darauf, daß ein Freund mir beim Umzug helfen wird, indem er den Transportwagen fährt. Es sieht so aus, als ob hier von Vertrauen nicht die Rede sein kann, solange ich nicht davon ausgehe, daß er auch wirklich Auto fahren kann; was ihn – den erwähnten Ansätzen zufolge – zum Teil vertrauenswürdig macht, ist die Tatsache, daß er über die Kompetenz verfügt, ein Fahrzeug zu steuern. Angenommen, ich vertraue darauf, daß meine Hausärztin mir ein geeignetes Medikament gegen meine Kopfschmerzen verschreiben wird: Auch hier kann nicht die Rede davon sein, daß ich ihr vertraute, wenn ich nicht zumindest implizit der Auffassung bin, daß sie als jemand, der Medizin studiert hat, über die Kompetenz verfügt, mir das richtige Medikament zu verschreiben; der ganze Witz von Zulassungen für praktizierende Ärzte besteht darin, in dieser basalen Hinsicht Vertrauenswürdigkeit sicherzustellen. Angenommen, ich vertraute darauf, daß mein Nachbar sich, während ich in den

<sup>16</sup> Baier (1986), 259, Hervorh. C. B.

<sup>17</sup> Jones (1996), 6–7, erste Hervorh. C. B. Obwohl sich Jones' Position in den letzten Jahren zum Teil massiv verändert hat, hält sie in ihren jüngsten Publikationen an Kompetenz als notwendiger Bedingung für Vertrauenswürdigkeit fest; vgl. dies. (2012), 70–71.

<sup>18</sup> Hardin (1999), 28, Hervorh. C. B.

<sup>19</sup> McLeod (2014), Hervorh. C. B.

Ferien bin, um meine Pflanzen kümmern wird: Auch hier muß ich, um vertrauen zu können, die Überzeugung haben (oder optimistisch sein), daß mein Nachbar schon wissen wird, was er tut, wenn er meine Pflanzen gießt.

Obwohl man nicht ernsthaft bestreiten kann, daß die in den Beispielfällen skizzierten Zusammenhänge tatsächlich vorliegen, denke ich, daß es problematisch ist, sie im Sinne von (KA) direkt in eine Analyse der Begriffe des Vertrauens und der Vertrauenswürdigkeit zu überführen. In einem ersten Schritt lassen sich meine Bedenken folgendermaßen formulieren: Die Frage, ob ich einer Person auf angemessene Weise *vertraue* bzw. ob eine andere Person *vertrauenswürdig* ist, kann nicht lediglich an Kompetenzen hängen, die in diesem Sinne *technisch* sind, daß sie im Prinzip unabhängig von mir und dem Verhältnis, in dem ich zu dieser Person stehe, zu denken sind. Für Vertrauen scheint es notwendig zu sein, daß die Person, die vertraut, und die Person, der vertraut wird, sich auf eine spezifische Weise aufeinander beziehen. Die in (KA) ausgedrückte Bedingung kommt allerdings völlig ohne eine solche Komponente aus.<sup>20</sup> Ob mein Freund Auto fahren kann, meine Hausärztin den Überblick über Kopfschmerzmedikamente hat oder mein Nachbar über einen grünen Daumen verfügt, hat zunächst gar nichts *mit mir*, der vertrauenden oder potentiell vertrauenden Person, zu tun.

Ein mit diesem Unbehagen zusammenhängendes Argument gegen (KA) greift den Unterschied zwischen Vertrauen und bloßem Sich-Verlassen wieder auf: Wenn es sich bei den Kompetenzen, auf die (KA) Bezug nimmt, um lediglich technische Kompetenzen handelt, dann scheint (KA) eher eine notwendige Bedingung für bloße Verlässlichkeit als für Vertrauenswürdigkeit zu formulieren. Ich kann mich nicht darauf verlassen, daß ich mit meinem Computer ein Paper fertigschreiben kann, wenn ich nicht implizit annehme, daß mein Computer eine Tastatur hat, sich hochfahren lassen wird, daß ein Betriebssystem und ein Textverarbeitungsprogramm darauf installiert sind usw. In dem Fall, in dem ich zu Recht davon ausgehe, daß mein Computer über diese ‚Kompetenzen‘ verfügt, werde ich mich zwar darauf verlassen können, daß ich mit dem Computer einen Text fertigstellen kann, aber es wäre abwegig davon zu reden, daß ich meinem Computer in dieser Hinsicht vertraute. Strukturell unterscheidet sich dieses Beispiel nicht wesentlich von den drei zuletzt skizzierten Beispielen. Das sieht man auch daran, daß die Reaktionen auf ein Scheitern in allen vier Fällen ähnlich ausfallen werden: Sollte mein Computer sich, nachdem ich mich endlich dazu durchgerungen habe, mit dem Schreiben zu beginnen, gar nicht erst hochfahren lassen, werde ich enttäuscht oder wütend auf mich selber sein, aber ich werde

---

<sup>20</sup> In der Terminologie von Jones müßte man hier von ‚starker Relationalität‘ sprechen; vgl. Jones (2010), 73.

mich nicht von dem Computer *betrogen fühlen*; auf ganz ähnliche Weise werde ich mich zunächst nicht betrogen fühlen, wenn ich lediglich feststellen sollte, daß mein Freund gar nicht Auto fahren kann, die medizinischen Kenntnisse meiner Hausärztin auf dem Stand der 50er Jahre steckengeblieben sind oder mein Nachbar der Auffassung ist, daß man Pflanzen nur einmal in der Woche gießen müsse, dann aber ordentlich.

Es scheint also, daß technische Kompetenzen auf Seiten der Person, der vertraut wird, für sich genommen keine für Vertrauen oder Vertrauenswürdigkeit konstitutive Bedingung darstellen: Wenn wir uns auf die inkompetente Vertrauensperson verlassen, dann machen wir zwar einen Fehler, aber es handelt sich dabei nicht um den Fehler, auf unangemessene Weise *vertraut* zu haben; der Fehler besteht vielmehr darin, daß wir uns auf die Person, der wir vertrauen, verlassen haben, ohne ihre Inkompetenz angemessen zu berücksichtigen. Aber heißt einer Person zu vertrauen nicht gerade, daß man sich auf sie verläßt? Warum – so der Einwand gegen die von mir formulierte These – sollte es einen Unterschied zwischen dem Vertrauen, das ich in eine Person setze, und der Tatsache, daß ich mich auf sie verlasse, geben? Das liegt eben daran, daß wir uns an dieser Stelle in dem theoretischen Rahmen von nicht-voluntaristischen Ansätzen bewegen und damit die These, daß Vertrauen eine Unterklasse von Sich-Verlassen darstellt, bereits hinter uns gelassen haben. Wenn Vertrauen ein Sich-Verlassen wäre, zu dem man sich entscheiden könnte, dann wäre Kompetenz tatsächlich ein guter Kandidat für eine notwendige Bedingung für Vertrauen. Dies sieht man auch daran, daß eine Version von (KA) eine wichtige Rolle im Rahmen voluntaristischer Ansätze spielt: Es handelt sich hier um die oben angesprochene Minimalbedingung, nach der die Person, der man vertraut, in der Lage sein muß, zu tun, was man von ihr erwartet. Eine Kompetenzbedingung – so das Resultat der bisherigen Überlegungen – hat ihren Platz in voluntaristischen Vertrauentheorien; in nicht-voluntaristischen Theorien – und das heißt auch: als Bestandteil eines Kriteriums für Vertrauenswürdigkeit – ist sie fehl am Platze.

## 4 Vertrauenswürdigkeit ohne Kompetenz

Anhand der zuletzt in den Blick genommenen Beispiele läßt sich zumindest in den Grundzügen ablesen, worin Vertrauenswürdigkeit stattdessen bestehen könnte. Sollte z. B. der Freund, der versprochen hat, mir beim Umzug zu helfen, indem er den Transporter fährt, sich als jemand herausstellen, der gar nicht Auto fahren kann und genau weiß, dass er gar nicht Auto fahren kann, ließe sich seine Ankündigung, als Fahrer zu helfen, nicht anders als als eine

implizite Lüge über seine Kompetenzen als Autofahrer interpretieren. In dieser Lesart hätte ich durchaus Grund, mich betrogen zu fühlen, aber dieser Grund bestünde nicht darin, daß der Freund bestimmte Kompetenzen hat oder nicht hat, sondern darin, daß er mich auf eine bestimmte Weise missachtet oder nicht ernst genommen hat, indem er mich täuschte. Sollte ich herausfinden, daß meine Ärztin mir ein bestimmtes Medikament verschrieben hat, weil sie dafür Geld von einem pharmazeutischen Unternehmen bekommt, oder daß mein Nachbar versprochen hat, meine Planzen zu gießen, weil er sich ungestört in meiner Wohnung umschaun möchte, hätte ich ebenfalls Grund, einen Vertrauensbruch zu beklagen und berechtigte Zweifel an der Vertrauenswürdigkeit der Ärztin bzw. des Nachbars zu haben. Wie genau man diese Beobachtungen für eine Bestimmung von Kriterien für Vertrauenswürdigkeit fruchtbar machen müßte, ist eine Frage, der ich in diesem Rahmen nicht nachgehen kann. Sollten sie allerdings in die richtige Richtung weisen, läßt sich wenigstens in den Ansätzen erkennen, inwiefern Vertreter von nicht-voluntaristischen Ansätzen etwas von der Strategie lernen können, die voluntaristische Ansätze einschlagen, um den Unterschied zwischen Vertrauen und bloßem Sich-Verlassen zu explizieren: Es werden hier nicht mehr bestimmte Eigenschaften in den Blick genommen, die eine Person unabhängig von anderen Personen haben kann und die sie als vertrauenswürdig auszeichnen; stattdessen geraten Aspekte in den Fokus, die mit der *Beziehung* zu tun haben, in der beide Vertrauenspartner zueinander stehen. Hier kann zwar von für Vertrauen notwendigen Kompetenzen die Rede sein, aber nur im Sinne von interpersonellen Kompetenzen, die von den in (KA) gemeinten Kompetenzen zu unterscheiden sind.

An dieser Stelle könnte eingewendet werden, daß meine Kritik an nicht-voluntaristischen Ansätzen – selbst wenn sie zutreffend sein sollte – nicht besonders weitreichend oder interessant ist. Immerhin formuliert (KA) im Rahmen solcher Ansätze nur *eine* notwendige Bedingung. Welche Konsequenzen sollte es schon haben, wenn man diese Bedingung einfach wegfallen läßt oder sie zu einer Ermöglichungsbedingung abschwächt? Ein wichtiger Kontext, in dem ein Verzicht auf (KA) als Bestandteil der Analyse von Vertrauenswürdigkeit eine Rolle spielt, ist der Kontext unserer sozialen Interaktionen mit Experten bzw. Personen, denen wir bestimmte Befugnisse übertragen haben. Es stellt sich hier die Frage, ob wir überhaupt Grund dazu haben, von der Vertrauenswürdigkeit von z. B. Ärztinnen<sup>21</sup> oder Regierungsmitgliedern<sup>22</sup> auszugehen, oder ob es nicht angemessener wäre, darauf zu verzichten, Vertrauen in sie zu legen. Ein wich-

---

21 Vgl. die Diskussion in O'Neill (2002).

22 Vgl. Hardin (1999), hier v. a. 33 ff.

tiges Problem scheint eben damit zu tun zu haben, daß wir in einer epistemisch sehr schlechten Position sind, die Kompetenzen dieser Personen einzuschätzen. Wenn meine Kritik an (KA) zutrifft, dann stellt das zwar ein Problem im Hinblick auf die Verlässlichkeit unserer Volksvertreter oder Ärzte dar, ist aber zunächst von der Frage nach Vertrauen zu trennen.

Was sollte es aber wiederum heißen, daß etwas problematisch für Verlässlichkeit, aber weniger problematisch für die Vertrauenswürdigkeit von Personen ist? Ist es nicht so, daß wir Vertrauenswürdigkeit an Personen schätzen, weil wir uns auf sie verlassen wollen? Es stimmt zwar, daß in der Regel Vertrauenswürdigkeit und Verlässlichkeit ‚gemeinsam auftreten‘; daß es nicht unwichtig ist, beide begrifflich sauber auseinanderzuhalten, sieht man aber an Situationen, in denen die Erwartungen einer Person enttäuscht werden – in denen uns eine Ärztin aus Inkompetenz das falsche Medikament verschreibt oder Mitglieder einer Regierung sich aus mangelndem Wissen für eine zum Scheitern verurteilte politische Maßnahme entscheiden. Wäre (KA) eine für Vertrauenswürdigkeit notwendige Bedingung, müßten wir in solchen Situationen Vertrauen entziehen oder zumindest überdenken. Vertrauensbrüche sind aber in der Regel weitaus schwerer zu reparieren als Enttäuschungen, die zur Folge haben, daß wir uns auf Personen nicht mehr verlassen. Anders gesagt: Sollte meine Ärztin mir glaubhaft machen können, daß sie sich inzwischen über für die für mich geeigneten Medikamente informiert hat, wäre es nicht irrational, zu unserem ursprünglichen Verhältnis zurückzukehren. Hält man an (KA) als Bedingung für Vertrauenswürdigkeit fest, ist nicht zu sehen, warum das so einfach sein sollte, weil man dieser Interpretation zufolge immerhin einen Vertrauensbruch zu überwinden hätte.

Die meiner Ansicht nach systematisch wichtigste Konsequenz aus der Ablehnung von (KA) besteht darin, daß durch diese Ablehnung ein Argument für eine weitere in der Debatte um den Begriff des Vertrauens vorherrschende Annahme entkräftet wird. Es handelt sich dabei um die folgende Relativitätsannahme:

(RA) Wenn eine Person A einer Person B vertraut, dann heißt das immer, daß A darauf vertraut, daß B etwas Bestimmtes tun (oder unterlassen) wird. Vertrauen ist ein dreistelliges Prädikat, und seine Zuschreibungen müssen immer relativ zu einer bestimmten Handlungserwartung spezifiziert werden. Zuschreibungen von Vertrauen, die keine solche Spezifikation enthalten, sind unvollständig.

In der einen oder anderen Form ist (RA) Bestandteil aller Ansätze in der Debatte um den Begriff des Vertrauens. Wer (RA) bestreiten möchte, vertritt die These, daß es vollständige Vertrauenszuschreibungen gibt, die nicht relativ zu einer bestimmten Handlungserwartung spezifiziert sind. Im Rahmen einer solchen

Zuschreibung wird behauptet, daß eine Person einer anderen Person *simpliciter* vertraut, d. h., es liegt hier keine bestimmte Hinsicht vor, in der eine Person einer anderen Person vertraut, und sie vertraut auch nicht darauf, daß die andere Person etwas Bestimmtes tun wird. Vertraut eine Person einer anderen Person *auf diese Weise*, dann läßt sich sagen, daß sie ihr – bis auf weiteres – in jeder Hinsicht vertraut, oder zumindest keine besondere Hinsicht spezifizieren kann, in der sie ihr vertraut oder nicht vertraut: Sie steht zu ihr in einer Vertrauensbeziehung. Versteht man Vertrauen auf diese Weise als ein zweistelliges Prädikat, sieht man sich relativ schnell mit dem folgenden einfachen Argument konfrontiert: Man kann einer anderen Person nicht in jeder Hinsicht vertrauen, weil Vertrauen notwendig eine Kompetenzeinschätzung impliziert und es eine triviale Wahrheit darstellt, daß es keine Person gibt, die in jeder Hinsicht kompetent ist. Dieses Argument ist so bestechend, daß jedes Projekt, das Vertrauen als ein zweistelliges Prädikat zu interpretieren versucht, bizarr und von Beginn an zum Scheitern verurteilt wirkt. Es ist allerdings nur in dem Ausmaß bestechend, in dem man bereit ist, (KA) als seine Prämisse zu akzeptieren. Ich hoffe, zumindest ansatzweise plausibel gemacht zu haben, daß es dafür keine guten Gründe gibt.<sup>23</sup>

## Literatur

- Baier, A. (1986), Trust and Antitrust, in: *Ethics* 96, 231–260.  
 Becker, L. C. (1996), Trust as Non-Cognitive Security about Motives, in: *Ethics* 107, 43–61.  
 Faulkner, P. (2007), On Telling and Trusting, in: *Mind* 116, 875–902.  
 Faulkner, P. (2011), *Knowledge On Trust*, Oxford.  
 Hardin, R. (1999), Do we want trust in government?, in: Warren, M. E. (Hg.), *Democracy and Trust*, Cambridge, 22–41.  
 Hardin, R. (2002), *Trust and Trustworthiness*, New York.  
 Hawley, K. (2012), Trust, Distrust and Commitment, in: *Noûs* 48.1, 1–20.  
 Hieronymi, P. (2008), The Reasons of Trust, in: *Australasian Journal of Philosophy* 86, 213–236.  
 Holton, R. (1994), Deciding to trust, coming to believe, in: *Australasian Journal of Philosophy* 72, 63–76.  
 Humberston, L. (1992), Direction of Fit, in: *Mind* 101, 59–83.  
 Jones, K. (1996), Trust as an Affective Attitude, in: *Ethics* 107.1, 4–25.  
 Jones, K. (2010), Counting On One Another, in: Grøn, A., u. Welz, C. (Hg.), *Trust, Sociality, Selfhood*, Tübingen, 67–82.  
 Jones, K. (2012), Trustworthiness, in: *Ethics* 123, 61–85.

---

<sup>23</sup> Für hilfreiche Kommentare zu früheren Fassungen des Textes danke ich Holger Baumann, Monika Betzler, Susanne Boshammer, Michael Kühler und den Teilnehmerinnen und Teilnehmern des Kolloquiums für Praktische Philosophie in Bern.



McLeod, C. (2002), *Self-Trust and Reproductive Autonomy*, Cambridge, Mass.

McLeod, C. (2014), Trust, in: Zalta, E. N. (Hg.), *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), URL: <http://plato.stanford.edu/archives/sum2014/entries/trust> (3.12.2015).

O'Neill, O. (2002), *Autonomy and Trust in Bioethics*, Cambridge.

Platts, M. (1979), *Ways of Meaning*, London.

Strawson, P. (1974), *Freedom and Resentment* [1962], in: ders., *Freedom and Resentment. And other Essays*, New York, 1–28.